Lista 3 - Econometria II

Guilhermo Argentieri

June 2025

Questão 1

(a)

A função de verossimilhança é dada por:

$$L(\theta, x_0|x) = \prod_{i=1}^{n} \theta x_0^{\theta} x_i^{-\theta-1} = \theta^n x_0^{n\theta} \left(\prod_{i=1}^{n} x_i\right)^{-\theta-1}$$

Aplicando o logaritmo natural, obtemos a seguinte função log-verossimilhança:

$$ln(L(\theta, x_0|x)) = n\ ln(\theta) + n\theta\ ln(x_0) - (\theta+1)\sum_{i=1}^{n} ln(x_i)$$

Note que a função de verossimilhança é crescente em x_0, mas deve cumprir $x_i > x_0 \forall i = 1, .., n$. Logo, o valor máximo de x_0 que satisfaz essa condição é:

$$\hat{x_0} = min(x_1, ..., x_n)$$

Agora, para achar o estimador para θ, vamos derivar a log-verossimilhança em relação a θ e igualar a 0:

$$\begin{aligned}
\frac{\partial \ln L(\theta)}{\partial \theta} &= \frac{n}{\theta} + n \ln x_0 - \sum_{i=1}^{n} \ln x_i = 0 \\
\frac{n}{\hat{\theta}} &= \sum_{i=1}^{n} \ln x_i - n \ln x_0 \\
\frac{n}{\hat{\theta}} &= \sum_{i=1}^{n} \ln\left(\frac{x_i}{x_0}\right) \\
\therefore \hat{\theta}_{MV} &= \frac{n}{\sum_{i=1}^{n}[\ln(x_i) - \ln(\hat{x_0})]}
\end{aligned}$$

Verificando a segunda derivada:

$$\frac{\partial^2 ln(L(\theta|x))}{\partial \theta^2} = -\frac{n}{\theta^2} < 0$$

Como a segunda derivada é negativa, então o ponto crítico é um máximo.

(b)

A função de verossimilhança é dada por:

$$L(\theta|x) = \prod_{i=1}^{n} \frac{1}{\theta}\mathbf{1}_{[0,\theta]}(x_i) = \theta^{-n} \prod_{i=1}^{n} \mathbf{1}_{[0,\theta]}(x_i)$$

Note que:

$$L(\theta|x) = \begin{cases} \theta^{-n}, \text{ se } \ max(x_1, ..., x_n) \leq \theta \\ 0, \text{caso contrário} \end{cases}$$

θ^{-n} é uma função decrescente em θ, logo, $L(\theta|x)$ é maximizada quando atinge seu menor valor que satisfaz $\theta \geq max(x_1, ..., x_n)$. Portanto, o estimador de máxima verossimilhança é dado por:

$$\hat{\theta}_{MV} = max(x_1, ..., x_n)$$

Obtendo a CDF de $\hat{\theta}_{MV}$:

$$F_{\hat{\theta}}(t) = P[max(x_1, ..., x_n) \leq t] = P(x_1 \leq t, x_2 \leq t, ..., x_n \leq t)$$

Note que:

$$P(x_i \leq t) = \begin{cases} 0, \text{ se } \ t < 0 \\ \frac{t}{\theta}, \text{ se } \ t \in [0, \theta] \\ 1, \text{ se } \ t > \theta \end{cases}$$

Então, para $t \in [0, \theta]$, temos:

$$F_{\hat{\theta}}(t) = \left(\frac{t}{\theta}\right)^n$$

Derivando em relação a t, obtemos a densidade:

$$f_{\hat{\theta}}(t) = \frac{dF_{\hat{\theta}}(t)}{dt} = n\frac{t^{n-1}}{\theta^n}$$

Com isso, calculamos a esperança de $\hat{\theta}_{MV}$:

$$\begin{aligned} \mathbb{E}[\hat{\theta}_{MV}] &= \int_0^\theta t \cdot n\frac{t^{n-1}}{\theta^n}\, dt \\ &= \frac{n}{\theta^n} \int_0^\theta t^n\, dt \\ &= \frac{n}{\theta^n} \left[\frac{t^{n+1}}{n+1}\right]_0^\theta \\ &= \frac{n}{\theta^n} \frac{\theta^{n+1}}{n+1} \\ &= \theta\frac{n}{n+1} \end{aligned}$$

Note que $\mathbb{E}[\hat{\theta}_{MV}] \neq \theta$, portanto, o estimador é viesado.

(c)

A função de verossimilhança é dada por:

$$L(\beta_1, \beta_2|x) = \prod_{i=1}^{n} \frac{1}{\beta_2 - \beta_1} \mathbf{1}_{[\beta_1,\beta_2]}(x_i) = (\beta_2 - \beta_1)^{-n} \prod_{i=1}^{n} \mathbf{1}_{[\beta_1,\beta_2]}(x_i)$$

Para que $L(\beta_1, \beta_2|x)$, é necessário $\beta_1 \leq min(x_1, ..., x_n)$ e $\beta_2 \geq max(x_1, ..., x_n)$, Note que $(\beta_2 - \beta_1)^{-n}$ é decrescente em $(\beta_2 - \beta_1)$, logo, $L(\beta_1, \beta_2|x)$ é maximizada quando $(\beta_2 - \beta_1)$ atinge seu menor valor. Portanto,

$$\hat{\beta}_{1MV} = min(x_1, ..., x_n)$$
$$\hat{\beta}_{2MV} = max(x_1, ..., x_n)$$

Questão 4

(a)

A função de verossimilhança é dada por:

$$L(\beta) = \prod_{i=1}^{n} \frac{e^{-e^{x_i\beta}} e^{x_i\beta y_i}}{y_i!}$$

Aplicando o log, obtemos:

$$\ell(\beta) = ln(L(\beta)) = \sum_{i=1}^{n} (e^{-x_i\beta} + \beta x_i y_i - y_i!)$$

(b)

Derivando e igualando a zero:

$$\frac{\partial \ell(\beta)}{\partial \beta} = \sum_{i=1}^{n} (x_i y_i - x_i e^{x_i\beta}) = 0$$

Note que não há solução analítica fechada para o estimador.Portanto, a CPO que caracteriza o estimador é:

$$\sum_{i=1}^{n} x_i(y_i - e^{x_i\hat{\beta}_{MV}}) = 0$$

(c)

Avaliando a segunda derivada da função log-verossimilhança:

$$\frac{\partial^2 \ell(\beta)}{\partial \beta^2} = \sum_{i=1}^{n} (-x_i^2 e^{x_i\beta}) < 0 \quad \forall x_i, \beta$$

Dado que o estimador de máxima verossimilhança satisfaz a CPO e que a segunda derivada da função log-verossimilhança é sempre negativa, a função é estritamente côncava, o que garante que $\hat{\beta}_{MV}$ maximiza a função de máxima verossimilhança.

(d)

A matriz de informação de Fisher é dada por:

$$I(\beta) = -\mathbb{E}\left[\frac{\partial^2 \ell(\beta)}{\partial\beta\beta^T}\right]$$

Do exercício anterior, teremos:

$$I(\beta) = -\mathbb{E}\left[\sum_{i=1}^{n}(-x_i^2 e^{x_i\beta})\right] = \sum_{i=1}^{n} x_i x_i^T e^{x_i\beta}$$

Portanto, sob as condições de regularidade, temos que:

$$\sqrt{n}(\hat{\beta}_{MV} - \beta_0) \xrightarrow{d} \mathcal{N}(0, I^{-1}(\beta_0))$$

Questão 7

(a)

A função de verossimilhança é dada por:

$$L(y_i; x_i; \beta) = \prod_{i=1}^{n} \frac{1}{\beta + x_i} e^{-\frac{y_i}{\beta + x_i}}$$

Aplicando o log, obtemos a seguinte log-verossimilhança:

$$\ell(\beta) = \sum_{i=1}^{n}\left(-\frac{y_i}{\beta + x_i} - log(\beta + x_i)\right) = -\sum_{i=1}^{n}\left(\frac{y_i}{\beta + x_i} + log(\beta + x_i)\right)$$

(b)

Derivando e igualando a zero:

$$\frac{\partial \ell(\beta)}{\partial \beta} = -\sum_{i=1}^{n}\left[-\frac{y_i}{(\beta + x_i)^2} + \frac{1}{\beta + x_i}\right] = 0 \Rightarrow \sum_{i=1}^{n}\left[\frac{(\beta + x_i) - y_i}{(\beta + x_i)^2}\right] = 0$$

Note que não há solução analítica fechada para β. Logo, a CPO que caracteriza o estimador é:

$$\sum_{i=1}^{n}\left[\frac{(\hat{\beta}_{MV} + x_i) - y_i}{(\hat{\beta}_{MV} + x_i)^2}\right] = 0$$

(c)

Para este item, será utilizado o método de Newton-Raphson, que utilizará o seguinte mecanismo para cada iteração:

$$\beta_{k+1} = \beta_k - \left[\frac{\partial^2 \ell(\beta)}{\partial \beta^2}\right]^{-1} \cdot \left[\frac{\partial \ell(\beta)}{\partial \beta}\right]$$

```r
library(readxl)

dados_q7 <- read_excel("data_PS3.xlsx", sheet = 'data-q5')
beta <- 1 #Valor inicial da estimação por Newton Ralphson
y <- as.numeric(dados_q7$Renda)
x <- as.numeric(dados_q7$Educ)


for(k in 1:10) {
#Calculando as derivadas
d1 <- sum( -1/(beta + x) + y/(beta + x)^2 )
d2 <- sum( 1/(beta + x)^2 - 2*y/(beta + x)^3 )

# update pelo passo de Newton
beta <- beta - d1/d2
cat("iteração", k, "-> beta =", beta, "\n")
}
```

```
## iteração 1 -> beta = 5.085934
## iteração 2 -> beta = 8.695518
## iteração 3 -> beta = 10.66269
## iteração 4 -> beta = 11.07962
## iteração 5 -> beta = 11.09468
## iteração 6 -> beta = 11.0947
## iteração 7 -> beta = 11.0947
## iteração 8 -> beta = 11.0947
## iteração 9 -> beta = 11.0947
## iteração 10 -> beta = 11.0947
```

Questão 8

##(c)

```r
library(AER)
library(gmm)

#Extraindo os dados
dados_q8 <- read_excel("data_PS3.xlsx", sheet = "data-q6")
y <- as.numeric(dados_q8$Y)
x <- as.numeric(dados_q8$X)
z <- as.numeric(dados_q8$Z)

#MQ2E (IV)
```

```
model_iv <- ivreg(y ~ x | z)
beta_iv <- coef(model_iv)

#GMM

gmm_iv <- gmm(y~x, x = z, wmatrix = 'optimal')
beta_gmm_iv <- coef(gmm_iv)

print(beta_iv)
```

```
## (Intercept)           x
##     9.36912     3.06382
```

```
print(beta_gmm_iv)
```

```
## (Intercept)           x
##     9.36912     3.06382
```

(d)

Temos o modelo linear $y_i = x_i'\beta + \varepsilon_i$, podendo ser reescrito como $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ se considerarmos $Y_i = y_i$, $x_i' = (1 \quad X_i)$ e $\epsilon_i = \varepsilon$. Ademais, consideraremos $z_i = (1 \quad Z_i \quad Z_i^2 \quad Z_i^3)$. A função momento para GMM com instrumentos z_i é dada por:

$$g(w_i, \beta) = z_i \cdot (Y_i - \beta_0 - \beta_1 X_i) = \begin{pmatrix} 1 \\ Z_i \\ Z_i^2 \\ Z_i^3 \end{pmatrix} \cdot (Y_i - \beta_0 - \beta_1 X_i) = \begin{bmatrix} (Y_i - \beta_0 - \beta_1 X_i) \\ Z_i(Y_i - \beta_0 - \beta_1 X_i) \\ Z_i^2(Y_i - \beta_0 - \beta_1 X_i) \\ Z_i^3(Y_i - \beta_0 - \beta_1 X_i) \end{bmatrix}$$

Portanto, as condições de momento são:

$$\mathbb{E}\left[\begin{pmatrix} 1 \\ Z_i \\ Z_i^2 \\ Z_i^3 \end{pmatrix} \cdot (Y_i - \beta_0 - \beta_1 X_i)\right] = 0$$

(e)

```
z_2 <- z^2
z_3 <- z^3

iv_8e <-  ivreg(y ~ x | z + z_2 + z_3)
gmm_8e <- gmm(y ~ x, ~ z + z_2 + z_3, wmatrix = 'optimal')

print(coef(iv_8e))
```

```
## (Intercept)           x
##    9.919521    2.067385
```

```r
print(coef(gmm_8e))
```

```
## (Intercept)           x
##    9.911935    2.072861
```

É possível que o estimador de GMM seja similar ao de 2SLS desde que definamos a matriz de pesos como $W_n = (Z'Z)^{-1}$

(f)

```r
library(tidyverse)
j_test <- (gmm_8e %>% summary)$stest
print(j_test)
```

```
##
## ##  J-Test: degrees of freedom is 2  ##
##
##                   J-test   P-value
## Test E(g)=0:    2.51088  0.28495
```

Dado que o p-valor é maior que o nível de significância, há evidências para rejeitarmos a hipótese nula, ou seja, temos indicativos de que as condições de ortogonalidade não são inválidas.

Questão 9

Substituindo $W_n = \Phi^{-1}$:

$$
\begin{aligned}
V &= (\Sigma'_{ZX}\Phi^{-1}\Sigma_{ZX})^{-1}(\Sigma'_{ZX}\underbrace{\Phi^{-1}\Phi}_{I}\Phi^{-1}\Sigma_{ZX})(\Sigma'_{ZX}\Phi^{-1}\Sigma_{ZX})^{-1} \\
\Rightarrow V &= \underbrace{(\Sigma'_{ZX}\Phi^{-1}\Sigma_{ZX})^{-1}(\Sigma'_{ZX}\Phi^{-1}\Sigma_{ZX})}_{I}(\Sigma'_{ZX}\Phi^{-1}\Sigma_{ZX})^{-1} \\
\therefore V &= (\Sigma'_{ZX}\Phi^{-1}\Sigma_{ZX})^{-1} = V_0
\end{aligned}
$$

(b)

Defina B como $B = \Phi^{-1}\Sigma_{ZX}(\Sigma'_{ZX}\Phi^{-1}\Sigma_{ZX})^{-1}$. Assim, podemos fazer a verificação:

$$
\begin{aligned}
B'\Phi B &= [\Phi^{-1}\Sigma_{ZX}(\Sigma'_{ZX}\Phi^{-1}\Sigma_{ZX})^{-1}]'\Phi[\Phi^{-1}\Sigma_{ZX}(\Sigma'_{ZX}\Phi^{-1}\Sigma_{ZX})^{-1}] \\
&= [(\Sigma'_{ZX}\Phi^{-1}\Sigma_{ZX})^{-1}]'\Sigma'_{ZX}\underbrace{\Phi^{-1}\Phi}_{I}\Phi^{-1}\Sigma_{ZX}(\Sigma'_{ZX}\Phi^{-1}\Sigma_{ZX})^{-1} \\
&= [(\Sigma'_{ZX}\Phi^{-1}\Sigma_{ZX})^{-1}]'\Sigma'_{ZX}\Phi^{-1}\Sigma_{ZX}(\Sigma'_{ZX}\Phi^{-1}\Sigma_{ZX})^{-1} \\
&= (\Sigma'_{ZX}\Phi^{-1}\Sigma_{ZX})^{-1} = V_0
\end{aligned}
$$

Defina A como $A = W_n\Sigma_{ZX}(\Sigma'_{ZX}W_n\Sigma_{ZX})^{-1}$. Assim, podemos fazer a verificação:

$$
\begin{aligned}
A'\Phi A &= [W_n\Sigma_{ZX}(\Sigma'_{ZX}W_n\Sigma_{ZX})^{-1}]'\Phi W_n\Sigma_{ZX}(\Sigma'_{ZX}W_n\Sigma_{ZX})^{-1} \\
&= (\Sigma'_{ZX}W_n\Sigma_{ZX})^{-1}\Sigma'_{ZX}W_n\Phi W_n\Sigma_{ZX}(\Sigma'_{ZX}W_n\Sigma_{ZX})^{-1} = V
\end{aligned}
$$

(c)

$$\begin{aligned}
B'\Phi A &= [\Phi^{-1}\Sigma_{ZX}(\Sigma'_{ZX}\Phi^{-1}\Sigma_{ZX})^{-1}]'\Phi W_n\Sigma_{ZX}(\Sigma'_{ZX}W_n\Sigma_{ZX})^{-1}\\
&= [(\Sigma'_{ZX}\Phi^{-1}\Sigma_{ZX})^{-1}]'\Sigma'_{ZX}\underbrace{\Phi^{-1}\Phi}_{I}W_n\Sigma_{ZX}(\Sigma'_{ZX}W_n\Sigma_{ZX})^{-1}\\
&= [(\Sigma'_{ZX}\Phi^{-1}\Sigma_{ZX})^{-1}]'\Sigma'_{ZX}\underbrace{W_n}_{\Phi^{-1}}\Sigma_{ZX}(\Sigma'_{ZX}W_n\Sigma_{ZX})^{-1}\\
&= (\Sigma'_{ZX}\Phi^{-1}\Sigma_{ZX})^{-1} = B'\Phi B
\end{aligned}$$

Logo, podemos notar que:

$$B'\Phi A = B'\Phi B \Rightarrow B'\Phi A - B'\Phi B = 0 \Rightarrow B'\Phi(A-B) = 0$$

(d)

Sabemos que a matriz Φ é positiva definida, pois, caso não fosse, o estimador de GMM não estaria bem definido. Portanto, podemos escrever a matriz como $\Phi = \Phi^{\frac{1}{2}}\Phi^{\frac{1}{2}}$. Ademais, vamos definir $C := \Phi^{\frac{1}{2}}A$ e $D := \Phi^{\frac{1}{2}}B$. Com isso, sabemos que $V = C'C$ e $V_0 = D'D$, também sabemos do item anterior que $D'C = D'D = V_0$. Com isso, temos:

$$V - V_0 = C'C - D'D$$

Dado que $D'C = D'D$, podemos substituir isso na equação e manipular da seguinte forma:

$$V - V_0 = C'C - D'C + C'D - D'D = (C-D)'(C-D)$$

Dado que o produto de uma matriz pela sua transposta é uma matriz positiva semi-definida, concluímos que:

$$V - V_0 = (C-D)'(C-D) \therefore V - V_0 \geq 0$$

Questão 10

(a)

No sistema, há duas variáveis endógenas (X e Y) e apenas uma variável exógena (Z). Na equação 1, o número de variáveis exógenas excluídas (zero) é maior do que o número de variáveis endógenas no sistema menos um ($2-1=1$), o que viola o critério da ordem, logo, β_2 não é identificado. Já na segunda equação, o número de variáveis exógenas excluídas (1) é igual ao número de variáveis endógenas do sistema menos um ($2-1=1$), portanto, α_2 é identificado

(b)

É possível estimar α_2 usando Z como instrumento utilizando a condição de momento $\mathbb{E}[ZV] = 0$. Sendo seu análogo amostral dado por $\frac{1}{n}\sum_{i=1}^{n} Z_i'V_i$. além disso, note que, $\mathbb{E}[V_i] = 0 \Rightarrow \mathbb{E}[X_i - \alpha_1 - \alpha_2 Y_i] = 0$ com análogo amostral dado por $\frac{1}{n}\sum_{i=1}^{n}(X_i - \alpha_1 - \alpha_2 Y_i)$

Dado que existem dois parâmetros e dois momentos, o modelo está exatamente identificado. Os parâmetros podem ser estimados resolvendo o sistema:

$$
\begin{cases}
\sum_{i=1}^{n} Z_i(X_i - \alpha_1 - \alpha_2 Y_i) = 0 \\
\sum_{i=1}^{n} (X_i - \alpha_1 - \alpha_2 Y_i) = 0
\end{cases}
$$